# Online Supplement

Ng et al, A time for every purpose: using time-dependent sensitivity analysis to help manage and understand dynamic ecological systems, *The American Naturalist*.

## S1 Parameter values for the introductory model

In this section, we provide the parameter values of the introductory model Eqns. (3) and (4) used to illustrate the adjoint method. As a reminder, the model describes a population in a sink habitat that is currently maintained through immigration, but the habitat is being restored so eventually the population will become self-sustaining. We use the abbreviation PU for the arbitrary population unit, and VU for the arbitrary value unit.

- Unregulated per-capita birth rate: We choose $b = 1/\text{year}$.
- Per-capita loss rate: We want $\mu(t)$ to decrease as a sigmoid, so we choose

$$\mu(t) = \mu_0 + (\mu_1 - \mu_0)/(1 + e^{(t-t_0)/\tau}), \tag{S1}$$

  where $\mu_0 = 1.5/\text{year}$ and $\mu_1 = 0.5/\text{year}$ are the pre- and post-restoration per-capita loss rates, $t_0 = 10$ years the time at the inflection point of the sigmoid, and $\tau = 2$ years a timescale that characterises the steepness of the sigmoid.
- Coefficient for intraspecific competition: We choose $a = 0.1/\text{PU}$
- Immigration rate: We choose $\sigma = 0.2$ PU/year.
- Per-capita rate of contribution to ecosystem service: We choose $w = 1$ VU/year/PU.
- Per-capital terminal payoff: In this example, any perturbation will eventually decay downstream, so it is possible to eliminate the effects of a finite time horizon if we choose $v$ such that it is equal to the ecosystem service contribution had the time horizon been extended indefinitely beyond $T$. To estimate this, we linearise Eqn. (3) about the post-restoration carrying capacity $K$, and find that any perturbation will decay exponentially at a rate $\mu_1 - b(1-2aK)$ and hence contribute a reward of $w/[\mu_1 - b(1-2aK)]$. Based on this reward, we choose $v = 1.74$ VU/PU.
- Initial conditions: We want $x(0)$ to be the steady-state population pre-restoration. Solving the equation $bx(0)(1-ax(0)) - \mu_0 x(0) + \sigma = 0$ gives us $x(0) = 0.37$ PU.

## S2 Incorporating perturbation costs into time-dependent sensitivities

Just like in optimal control theory, we now consider a manipulated system

$$\frac{dx(t)}{dt} = g(\vec{x}(t),u(t),t), \quad \vec{x}(0) = \vec{x}_0, \tag{S2}$$

where $u(t)$ quantifies the external manipulation. We also define

$$K \equiv \int_0^T c(\vec{x}(t),u(t),t)dt, \tag{S3}$$

the total cost of the manipulation, analogous to the total reward function $J$. If there is no manipulation, there is no manipulation cost, so we require that $c(\vec{x},0,t) = 0$ for any $\vec{x}$ and $t$. At the same time, we assume that the integrand $f(\vec{x}(t),t)$ of the total reward $J$ does not depend directly on $u(t)$.

We are interested in the effects of a small, brief manipulation at time $t^*$ on the net value $J-K$. More specifically, we consider $u = \epsilon h$, where $h$ is a narrow window function centered at time $t^*$, normalized such that $\int_0^T h(t)dt = 1$. Since $J$ is only indirectly affected by the manipulation through the effects on $\vec{x}(t)$, if we interpret $u$ as yet another parameter with an unperturbed value of 0, we can apply Eqn. (A10) from Appendix B, so

$$\Delta J \simeq \epsilon \sum_j \lambda_j(t^*) \left.\frac{\partial g_j(\vec{x}(t^*),u(t^*),t^*)}{\partial u}\right|_{u(t^*)=0} \tag{S4}$$

Meanwhile, since $c(\vec{x},0,t) = 0$ for any $\vec{x}$ and $t$, this is also true for its partial derivative in $\vec{x}$, so to order $\mathcal{O}(\epsilon)$, $\Delta K$ only comes from the direct dependence of $c$ on $u$. More specifically,

$$\Delta K = \int_0^T \left.\frac{\partial c(\vec{x}(t),u(t),t)}{\partial u}\right|_{u(t)=0} \epsilon h(t)dt \simeq \left.\frac{\partial c(\vec{x}(t^*),u(t^*),t^*)}{\partial u}\right|_{u(t^*)=0} \epsilon, \tag{S5}$$

where in the second step, we used the fact that $h$ is a normalized narrow window function centered at time $t^*$. Hence, the sensitivity to a small, brief manipulation at time $t^*$ is given by

$$\boxed{\lim_{\epsilon\to 0} \frac{\Delta J - \Delta K}{\epsilon} = \sum_j \lambda_j(t^*) \left.\frac{\partial g_j(\vec{x}(t^*),u(t^*),t^*)}{\partial u}\right|_{u(t^*)=0} - \left.\frac{\partial c(\vec{x}(t^*),u(t^*),t^*)}{\partial u}\right|_{u(t^*)=0}.} \tag{S6}$$

Note that unlike optimal control theory, we only need the linearized versions of the functions $g_j$ and $c$ about $u = 0$ and not their full functional forms in order to calculate the sensitivity.

## S3 Change of adjoint variables under a change of state variables

Let $\vec{x}$ be the original state variables, and $\vec{y}$ be the new state variables. For simplicity, assume that the transformation is invertible and also has no explicit time dependence, so we can write each new variable $y_i$ as a function $y_i(\vec{x})$ of the old variables, and each old variable as a function $x_i(\vec{y})$ of the new variables.

When taking partial derivatives, it is important to keep track of what other variables are being held constant. We will use the notation $(\frac{\partial}{\partial x_i})_x$ to mean holding all other $x_{j\neq i}$ constant. The old and new variables satisfy the dynamic equations

$$\frac{dx_i}{dt} = g_{x_i}(\vec{x}(t),t), \quad \frac{dy_i}{dt} = g_{y_i}(\vec{y}(t),t). \tag{S7}$$

Since the transformation does not contain any explicit time dependence, chain rule tells us that

$$\frac{dy_i}{dt} = \sum_j \left(\frac{\partial y_i}{\partial x_j}\right)_x \frac{dx_j}{dt} = \sum_j \left(\frac{\partial y_i}{\partial x_j}\right)_x g_{x_j}, \tag{S8}$$

so we have the relation and inverse relation

$$g_{y_i} = \sum_j \left(\frac{\partial y_i}{\partial x_j}\right)_x g_{x_j}, \quad g_{x_i} = \sum_j \left(\frac{\partial x_i}{\partial y_j}\right)_y g_{y_j} \tag{S9}$$

Let the reward function be

$$J = \int_0^T f(\vec{x}(t),t)dx + \Psi(\vec{x}(T)). \tag{S10}$$

The old adjoint variables satisfy the adjoint equations and terminal conditions

$$\frac{d\lambda_{x_i}}{dt} = -\left(\frac{\partial f}{\partial x_i}\right)_x - \sum_j \lambda_{x_j}\left(\frac{\partial g_{x_j}}{\partial x_i}\right)_x, \quad \lambda_{x_i}(T) = \left(\frac{\partial \Psi}{\partial x_i}\right)_x\Big|_{\vec{x}=\vec{x}(T)}, \tag{S11}$$

while the new adjoint variables satisfy

$$\frac{d\lambda_{y_i}}{dt} = -\left(\frac{\partial f}{\partial y_i}\right)_y - \sum_j \lambda_{y_j}\left(\frac{\partial g_{y_j}}{\partial y_i}\right)_y, \quad \lambda_{y_i}(T) = \left(\frac{\partial \Psi}{\partial y_i}\right)_y\Big|_{\vec{y}=\vec{y}(T)}. \tag{S12}$$

In the remainder of this section, we will prove the relation

$$\boxed{\lambda_{y_i} = \sum_j \left(\frac{\partial x_j}{\partial y_i}\right)_y \lambda_{x_j}.} \tag{S13}$$

First, we define

$$\lambda_i' \equiv \sum_j \left(\frac{\partial x_j}{\partial y_i}\right)_y \lambda_{x_j}. \tag{S14}$$

Our strategy is to show that $\lambda_i'$ satisfies the same adjoint equations and terminal conditions as $\lambda_{y_i}$, so we can then conclude that $\lambda_i' = \lambda_{y_i}$, hence proving the relation. Consider

$$\begin{aligned}
\frac{d\lambda_i'}{dt} &= \underbrace{\frac{d\left(\sum_j \left(\frac{\partial x_j}{\partial y_i}\right)_y \lambda_{x_j}\right)}{dt}}_{\text{definition of } \lambda_i'} = \underbrace{\sum_j \left(\frac{\partial x_j}{\partial y_i}\right)_y \frac{d\lambda_{x_j}}{dt} + \sum_j \lambda_{x_j} \frac{d\left(\frac{\partial x_j}{\partial y_i}\right)_y}{dt}}_{\text{from product rule}} \\
&= \sum_j \left(\frac{\partial x_j}{\partial y_i}\right)_y \underbrace{\left[-\left(\frac{\partial f}{\partial x_j}\right)_x - \sum_k \lambda_{x_k} \left(\frac{\partial g_{x_k}}{\partial x_j}\right)_x\right]}_{\text{from adjoint equations Eqn. (S11)}} + \sum_j \lambda_{x_j} \underbrace{\sum_k \overbrace{\frac{dy_k}{dt}}^{g_{y_k}} \left(\frac{\partial^2 x_j}{\partial y_i \partial y_k}\right)_y}_{\text{from chain rule}} \\
&= \underbrace{-\sum_j \left(\frac{\partial x_j}{\partial y_i}\right)_y \left(\frac{\partial f}{\partial x_j}\right)_x}_{\left(\frac{\partial f}{\partial y_i}\right)_y} \underbrace{-\sum_j \sum_k \left(\frac{\partial x_j}{\partial y_i}\right)_y \lambda_{x_k} \left(\frac{\partial g_{x_k}}{\partial x_j}\right)_x}_{(*)} + \sum_j \sum_k \lambda_{x_j} g_{y_k} \left(\frac{\partial^2 x_j}{\partial y_i \partial y_k}\right)_y.
\end{aligned} \tag{S15}$$

We will first simplify the term $(*)$ before returning to the equation. We have

$$\begin{aligned}
(*) &= \sum_j \sum_k \lambda_{x_k} \left(\frac{\partial x_j}{\partial y_i}\right)_y \left(\frac{\partial g_{x_k}}{\partial x_j}\right)_x = \sum_j \sum_k \lambda_{x_k} \left(\frac{\partial x_j}{\partial y_i}\right)_y \underbrace{\sum_m \left(\frac{\partial y_m}{\partial x_j}\right)_x \left(\frac{\partial g_{x_k}}{\partial y_m}\right)_y}_{\text{from chain rule}} \\
&= \sum_k \sum_m \lambda_{x_k} \underbrace{\sum_j \left(\frac{\partial x_j}{\partial y_i}\right)_y \left(\frac{\partial y_m}{\partial x_j}\right)_x}_{\delta_{i,m}} \left(\frac{\partial}{\partial y_m} \underbrace{\left[\sum_n \left(\frac{\partial x_k}{\partial y_n}\right)_y g_{y_n}\right]}_{\text{from Eqn. (S9)}}\right)_y \\
&= \sum_k \sum_n \lambda_{x_k} \underbrace{\sum_m \delta_{i,m} \left(\frac{\partial}{\partial y_m}\right)}_{\frac{\partial}{\partial y_i}} \left[\sum_n \left(\frac{\partial x_k}{\partial y_n}\right)_y g_{y_n}\right]_y \\
&= \sum_k \sum_n \lambda_{x_k} \underbrace{\left[\left(\frac{\partial x_k}{\partial y_n}\right)_y \left(\frac{\partial g_{y_n}}{\partial y_i}\right)_y + \left(\frac{\partial^2 x_k}{\partial y_n \partial y_i}\right)_y g_{y_n}\right]}_{\text{from product rule}}.
\end{aligned} \tag{S16}$$

Now we replace the dummy variables $k$ and $n$ in $(*)$ by $j$ and $k$ respectively, and plug it back into

Eqn. (S15). We get

$$\begin{aligned}\frac{d\lambda_i'}{dt} &= \left(\frac{\partial f}{\partial y_i}\right)_y - \sum_j\sum_k \lambda_{x_j}\left(\frac{\partial x_j}{\partial y_k}\right)_y\left(\frac{\partial g_{y_k}}{\partial y_i}\right)_y \underbrace{- \sum_j\sum_k \lambda_{x_j}\left(\frac{\partial^2 x_j}{\partial y_k \partial y_i}\right)_y g_{y_k} + \sum_j\sum_k \lambda_{x_j} g_{y_k}\left(\frac{\partial^2 x_j}{\partial y_i \partial y_k}\right)_y}_{\text{cancels}} \\ &= \left(\frac{\partial f}{\partial y_i}\right)_y - \sum_k \lambda_k'\left(\frac{\partial g_{y_k}}{\partial y_i}\right)_y .\end{aligned} \tag{S17}$$

Comparing Eqn. (S17) to Eqn. (S12), we see that $\lambda_i'$ does indeed satisfy the same adjoint equations in Eqn. (S12) as $\lambda_{y_i}$. All that is left is to show that $\lambda_i'$ also satisfy the same terminal conditions in Eqn. (S12). Consider

$$\lambda_i'(T) = \sum_j \left(\frac{\partial x_j}{\partial y_i}\right)_y \lambda_{x_j}(T) = \sum_j \left(\frac{\partial x_j}{\partial y_i}\right)_y \underbrace{\left(\frac{\partial \Psi}{\partial x_j}\right)_x\Big|_{\vec{x}=\vec{x}(T)}}_{\text{from Eqn. (S11)}} = \left(\frac{\partial \Psi}{\partial y_i}\right)_y\Big|_{\vec{y}=\vec{y}(T)}, \tag{S18}$$

hence completing the proof.

More elegant proofs probably exist from optimal control theory, but this version is the most straightforward.

# S4 Parameter values for Example 1: Disease spillover into multi-species sink communities

As mentioned in the main text, the parameter values have been chosen to best illustrate the qualitative features of interest. We explain the choices in more details below.

- Disease-free mortality ($\mu_j$): For simplicity, we assume that all species have the same $\mu_j$. Without loss of generality, we choose the units of time so that one unit corresponds to one lifespan, so $\mu_j = 1$ for all $j$.
- Unregulated per-capita birth rate ($B_j$): For the species of concern, we want there to be a substantial population decline despite the low infection prevalence (especially if the disease reaches the species of concern from the exogenous source only after a long chain of transmission), so that control measures are necessary. Therefore, we choose $B_{j_C} = 1.02$ so that it is only very slightly above $\mu_{j_C}$.

  For all other species, as explained in the main text, culling an intermediate species too early in the season is ineffective since the population would have mostly recovered by the time the chain of infection reaches the species. To demonstrate this point clearly, we want $B_j \gg \mu_j$, so we choose $B_j = 5$.
- Intraspecific competition coefficient ($a_j$) or carrying capacity ($K_j$): We can specify either $a_j$ or $K_j$ since they are related by $K_j = (1 - \mu_j/B_j)/a_j$. For simplicity, we assume that all species have the same $K_j$, and without loss of generality, we choose the units of population size so that $K_j = 1$ for all $j$. This means that $a_j = 0.8$ for all species, except the species of concern, where $a_{j_C} \simeq 0.02$. In other words, the

large carrying capacity in the species of concern despite the low birth rate is due to low intraspecific competition.

Alternatively, we could have chosen the same competition coefficient $a_j = 0.8$ for all $j$, in which case all species will have $K_j = 1$ except for the species of concern, where $K_{j_C} \simeq 0.02$, i.e. a low carrying capacity. We find that most qualitative features observed in the two networks are still present under this alternative scenario.

- Disease-induced mortality ($\nu_j$): We want a large disease-induced mortality in the species of concern, so we choose $\nu_{j_C} = 5$. In contrast, for all other species, we choose $\nu_j = 0$, so the disease has no impact on their populations.
- Recovery rate ($\gamma_j$): Again, for there to be a substantial population decline in the species of concern, we need a high per-capita rate of infection in the species of concern, even after a long chain of transmission, while still keeping $R_0 < 1$. Numerically, we find that this is easiest to achieve when all species have comparable infectious lifetimes $1/(\mu_j + \nu_j + \gamma_j)$. Since the species of concern already has a short infectious lifetime due to the large disease-induced mortality $\nu_{j_C}$, we set $\gamma_{j_C} = 0$. For all other species without disease-induced mortality, we choose $\gamma_j = 5$, so that they recover quickly from infection.
- Length of active season ($T$): Even though both networks were meant to be hypothetical, we designed them with pollinators in mind. Since the average lifespan of a bee is of order 20–30 days, we choose $T = 5$ so that the active season would correspond to a realistic period of 100–150 days.
- Coefficients in the reward function ($W_{S_{j_C}}$, $W_{I_{j_C}}$, $V_{S_{j_C}}$, $V_{I_{j_C}}$): Without loss of generality, we choose the units of value so that $W_{S_{j_C}} = 1$. We assume that infected individuals are just as capable of providing the ecosystem service, so $W_{I_{j_C}} = 1$ as well. (One possible scenario is that most infected individuals in the species of concern start off as asymptomatic carriers, but quickly die once the symptoms set in. Therefore, the fecundity of infected individuals as well as the ecosystem service they provide remain unaffected before they die.) For the terminal payoffs, we arbitrarily choose $V_{S_{j_C}} = V_{I_{j_C}} = 1$. We find that most qualitative features observed in the networks are still present under other choices of $W_{I_{j_C}}$, $V_{S_{j_C}}$ and $V_{I_{j_C}}$.
- Transmission coefficients ($b_{j,k}$): We parametrize $b_{j,k}$ according to the network structure and then rescale them so that the dominant eigenvalue of the next-generation matrix is $R_0$. Below, we present the values of $b_{j,k}$ *before* rescaling.
  - Network 1: We take the $c \to \infty$ limit of the trait-matching model, which gives

$$\boldsymbol{B} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}. \tag{S19}$$

– Network 2: We first define resource utilization $r_{j,k}$ as the relative frequency an individual of species $k$ chooses to utilize resource type $j$. As explained in the main text, there are two resource types, and bridge species 3 (the species of concern) is less specialized, so we choose

$$\boldsymbol{r} = \begin{pmatrix} 1 & 1 & 0.2 & 0 & 0 \\ 0 & 0 & 0.8 & 1 & 1 \end{pmatrix} \tag{S20}$$

We then assume that $\boldsymbol{B}$ is given by $\boldsymbol{B} = \boldsymbol{r}^T \boldsymbol{r}$. To enhance intraspecific transmission in species 5, we also double the value of $b_{5,5}$.

- Basic reproduction number ($R_0$): We choose $R_0 = 0.9$ for Network 1, and $R_0 = 0.95$ for Network 2.
- Spillover coefficient ($\sigma_j$): In both networks, only the first species receive exogenous spillover. We choose $\sigma_1 = 0.2$ for both networks.
- Initial conditions ($S_j(0)$, $I_j(0)$): We choose $S_j(0) = K_j$ and $I_j(0) = 0$ for all $j$. In other words, we assume that each species starts the current season disease-free at the carrying capacity. This is mainly for simplicity, so that the transient dynamics mostly reflect disease transmission and not population growth.

# S5 More details on Example 2: Leopard frogs as reservoirs of the amphibian chytrid fungus

## *S5.1 Functional forms and parameter values*

The load-dependent functions $\ell(x)$, $G_0(x)$ and $G(x'|x)$ are assumed to take the form

$$\begin{aligned} \ell(x) &= 1 - \Phi(x|\mu_l, \sigma_l), \\ G_0(x') &= \phi(x'|a(t), \sigma_0), \\ G(x'|x) &= \phi(x'|a(t) + bx, \sigma_0). \end{aligned} \tag{S21}$$

Here $\phi$ and $\Phi$ are the probability density and cumulative distribution functions of the normal distribution, with mean and standard deviation given by the two parameters after the vertical bars.

The temperature-dependent functions $a(T)$ and $s_Z(T)$ are assumed to take the form

$$\begin{aligned} a(T) &= a_0 + a_1(T - T_{\text{base}}), \\ s_Z(T) &= \frac{s_{Z,0}}{1 + e^{(T - T_Z)/\sigma_Z}}, \end{aligned} \tag{S22}$$

The temperature is assumed to vary sinusoidally across the year, and is given by

$$T(t) = T_{\min} + \frac{T_{\max} - T_{\min}}{2}\left[1 - \cos\left(\frac{2\pi t}{52}\right)\right], \tag{S23}$$

where $t$ here is in weeks, and it is assumed that one year has exactly 52 weeks.

Wilber et al. (2022) fitted separate *Bd* transmission models at four geographic locations (Louisiana, Tennessee, Pennsylvania, and Vermont), and at three possible values of the parameter $K$ controlling density dependence in recruitment: $e^{10}$ (low density), $e^{8}$ (medium density) and $e^{4}$ (high density). Most parameter values can be found in Table S2 from Wilber et al. (2022); we chose parameter values for Tennessee under the high-density assumption, as well as $s_I = 1$. Other parameter values that can only be found in the main text or in their scripts are: $T_{\min} = 4°\text{C}$, $T_{\max} = 27°\text{C}$, aquatic calendar days 30–150 (so $W(t) = 1$ for week numbers 5–21), and reproduction calendar day 90 (so $R(t) = 1$ for week number 13).

## *S5.2 Discretizing the IPM*

We discretize the IPM in Eqn. (20) into $m$ bins each of width $h$. The $i$th bin has midpoint $x_i$, lower and upper boundaries $\underline{x}_i$ and $\overline{x}_i$, and contains $I_i(t)$ infected individuals (so $I_i(t)$ approximates $I(x_i,t)h$). The discretized equations are then given by

$$\begin{aligned}
L(t+1) &= r'\frac{N(t)}{2}R(t) + L(t)s_L(1-m_L),\\
S(t+1) &= L(t)s_L m_L e^{-KN(t)} + S(t)s_0 e^{-\beta Z(t)W(t)} + s_0 s_I \sum_{i=1}^{m} \ell_i I_i(t),\\
I_i(t+1) &= S(t)s_0\left(1 - e^{-\beta Z(t)W(t)}\right)(G_0)_i + s_0 s_I \sum_{j=1}^{m}(1-\ell_j)G_{ij}I_j(t),\\
Z(t+1) &= \lambda W(t)\sum_{i=1}^{m} e^{x_i} I_i(t) + s_Z(t)Z(t) + \omega,
\end{aligned} \tag{S24}$$

where

$$\begin{aligned}
N(t) &= S(t) + \sum_{i=1}^{m} I_i(t),\\
\ell_i &= 1 - \Phi(x_i|\mu_l,\sigma_l),\\
(G_0)_i &= \Phi(\overline{x}_i|a(t),\sigma_0) - \Phi(\underline{x}_i|a(t),\sigma_0),\\
G_{ij} &= \Phi(\overline{x}_i|a(t)+bx_j,\sigma_0) - \Phi(\underline{x}_i|a(t)+bx_j,\sigma_0).
\end{aligned} \tag{S25}$$

## S5.3 *Deriving the adjoint equations*

To derive the adjoint equations, we first write down the Hamiltonian

$$\begin{aligned} H = {} & \lambda_L(t+1)\cdot\left[r'\frac{S(t)+\sum_{i=1}^{m}I_i(t)}{2}R(t)+L(t)s_L(1-m_L)\right] \\ & +\lambda_S(t+1)\cdot\left[L(t)s_L m_L e^{-KS(t)-K\sum_{i=1}^{m}I_i(t)}+S(t)s_0 e^{-\beta Z(t)W(t)}+s_0 s_I\sum_{i=1}^{m}\ell_i I_i(t)\right] \\ & +\sum_{i=1}^{m}\lambda_{I,i}(t+1)\cdot\left[S(t)s_0\left(1-e^{-\beta Z(t)W(t)}\right)(G_0)_i+s_0 s_I\sum_{j=1}^{m}(1-\ell_j)G_{ij}I_j(t)\right] \\ & +\lambda_Z(t+1)\cdot\left[\lambda W(t)\sum_{i=1}^{m}e^{x_i}I_i(t)+s_Z(t)Z(t)+\omega\right]-V(t)Z(t). \end{aligned} \tag{S26}$$

We then obtain the adjoint equations, Eqn. (22), by taking partial derivatives of the Hamiltonian $H$ according to Eqn. (11).

# S6 More details on Example 3: Population cycles in the pine looper and the larch budmoth

## *S6.1 Larch budmoth: Model details*

Johnson et al. (2004, 2006) proposed a tritrophic, spatially-explicit, discrete-time model, where budmoths and their parasitoids are located in patches of suitable habitats embedded within a larger landscape. In each patch, which we index by $i$ (maximum $n$), and at year $t$, the local densities of budmoths and parasitoids are represented by state variables $H(i,t)$ and $P(i,t)$, while the local plant quality is represented by the state variable $Q(i,t)$ with a maximum value of 1. The dynamics can be represented by the equations

$$\begin{aligned}
H(i,t+1) &= \sum_{j=1}^{n} \left\{ \underbrace{\frac{e^{-(d_{ij}/\alpha_H)^2}}{C_H}}_{\substack{\text{budmoth}\\ \text{dispersal}}} \; H_j \underbrace{\exp\left[ r_0 \left( 1 - e^{-Q(j,t)/\delta} - \frac{H(j,t)}{k} \right) \right]}_{\text{local budmoth growth}} \; \underbrace{\exp\left( -\frac{aP(j,t)}{1+awP(j,t)} \right)}_{\text{avoiding local parasitism}} \right\}, \\
P(i,t+1) &= \sum_{j=1}^{n} \left\{ \underbrace{\frac{e^{-(d_{ij}/\alpha_P)^2}}{C_P}}_{\substack{\text{parasitoid}\\ \text{dispersal}}} \; \underbrace{H_j \left[ 1 - \exp\left( -\frac{aP(j,t)}{1+awP(j,t)} \right) \right]}_{\text{local parasitism}} \right\}, \\
Q(i,t+1) &= \underbrace{(1-\beta)+\beta Q(i,t)}_{\text{local plant recovery}} - \underbrace{\frac{uH(i,t)}{v+H(i,t)}}_{\text{local herbivory}} .
\end{aligned} \tag{S27}$$

For dispersal, $d_{ij}$ is the distance between patches, and we assume a Gaussian kernel with dispersal parameters $\alpha_H$ and $\alpha_P$ for the budmoths and parasitoids; $C_H$ and $C_P$ are normalization constants. Before dispersal, we assume that the local budmoth and parasitoid densities change in accordance to the local dynamics. For the budmoth, $r_0$ is the maximum growth rate[2], $\delta$ is a scale parameter that determines how fast the growth rate approaches $r_0$ with increasing plant quality $Q(j,t)$, and $k$ is the budmoth carrying capacity in the limit of large $Q(j,t)$, so $1/k$ characterizes intraspecific competition. Local parasitism is described by a modified Nicholson-Bailey framework: the exponential describes the probability of a budmoth avoiding parasitism, and is parametrized by $a$ and $w$ representing the search efficiency of a parasitoid and the mutual interference between parasitoids. Finally, for local plant dynamics, $\beta$ represent the rate at which plant quality $Q(i,t)$ recovers towards 1, while $u$ and $v$ characterize the impact of budmoth herbivory on plant quality. We note that Johnson et al. (2004) also introduced an additional parameter that is meant to approximate the effects of demographic stochasticity, although it was omitted in Johnson et al. (2006); we chose to omit it as well.

[2]Or nearly so, since $Q(j,t)$ cannot exceed 1, so the maximum growth rate is really $r_0(1-e^{-1/\delta}) \simeq 0.989 r_0$ for the chosen value of $\delta = 0.22$.

Most parameter values can be found in Table 1 of Johnson et al. (2006), although note that the parameter labels $(r_0,K,A,W,A,C,D,\delta)$ should be corrected to $(r_0,k,a,w,\beta,u,v,\delta)$. Other parameter values that can only be found in the main text are: $\alpha_H = 10$ km and $\alpha_P = 5$ km. For the normalization constants $C_H$ and $C_P$, the authors stated that they were chosen such that the "total proportion of dispersal across suitable and unsuitable habitat sums to one". Therefore, we discretized the landscape into an arbitrarily large spatial grid of resolution $3 \times 3$ km (based on the patch dimensions in Johnson et al. (2004)), and assumed that the Gaussian kernel applied to any pair of grid cells, and not just grid cells assigned as suitable patches. We then obtained $C_H$ using

$$C_H = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} e^{(i^2+j^2)/(\alpha_H/(3\text{ km}))^2}, \tag{S28}$$

where $i$ and $j$ here are grid indices (not patch indices). A similar expression was used for $C_P$.

We wanted to replicate the scenario in Johnson et al. (2004, 2006) where patches near the center of the landscape had the highest connectivity. According to Johnson et al. (2004), "habitat configurations were created by assuming that the probability of a patch being suitable declined exponentially with the distance from the focal location". Therefore, we drew random samples from an exponential distribution with a mean of 5 grid units, applied a random sign, and rounded them to the nearest integer. Pairs of these integers were then used as grid indices for the suitable patches. We generated 500 unique patches this way.

Since we were only interested in the deterministic version of the model, we did not introduce random variations into $r_0$ for each patch and timestep as was done in Johnson et al. (2006). Also, even though we initialized the simulation the same way as Johnson et al. (2006), we ran the simulation for many time steps before the start of the time horizon, to allow any transients to die off.

## *S6.2 Larch budmoth: Objective function and adjoint equations*

A possible objective function is to maximize the plant quality over a time horizon from $t = 1$ to $T$, with weight $W(i,t)$ assigned to patch $i$ at time $t$, so

$$J = \sum_{t=1}^{T-1} \sum_{i=1}^{n} W(i,t)Q(i,t) + \sum_{i=1}^{n} W(i,t)Q(i,T).$$

We choose an arbitrary time horizon of $T = 200$ years, and we assigned equal weight to all patches, but more weight to more recent years, by having

$$W(i,t) = e^{-t/\tau},$$

where $\tau = 50$ years. Just as in the pine looper example, the decaying weights reduce the dependence of the time-dependent sensitivities on the time horizon, should the dynamics be quasiperiodic.

The Hamiltonian (which we denote by $\mathcal{H}$ to avoid confusion with the budmoth density) is given by

$$\begin{aligned}\mathcal{H} = &\sum_{i=1}^{n}\lambda_H(i,t+1)\sum_{j=1}^{n}\left\{\frac{e^{-(d_{ij}/\alpha_H)^2}}{C_H}H(j,t)\exp\left[r_0\left(1-e^{-Q(j,t)/\delta}-\frac{H(j,t)}{k}\right)\right]\exp\left(-\frac{aP(j,t)}{1+awP(j,t)}\right)\right\}\\ &+\sum_{i=1}^{n}\lambda_P(i,t+1)\sum_{j=1}^{n}\left\{\frac{e^{-(d_{ij}/\alpha_P)^2}}{C_P}H(j,t)\left[1-\exp\left(-\frac{aP(j,t)}{1+awP(j,t)}\right)\right]\right\}\\ &+\sum_{i=1}^{n}\lambda_Q(i,t+1)\left[(1-\beta)+\beta Q(i,t)-\frac{uH(i,t)}{v+H(i,t)}\right]\\ &+\sum_{i=1}^{n}W(i,t)Q(i,t),\end{aligned} \tag{S29}$$

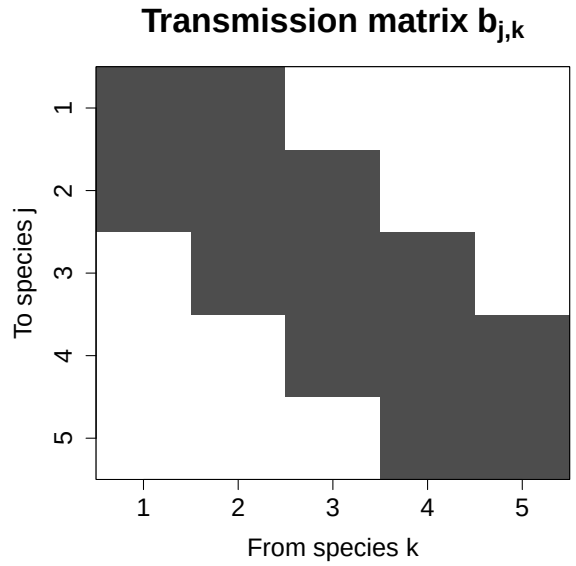
where the last term comes from the objective function. The adjoint equations are then given by

$$\begin{aligned}\lambda_H(i,t) = \frac{\partial\mathcal{H}}{\partial H(i,t)} = &\sum_{j=1}^{n}\lambda_H(j,t+1)\left\{\frac{e^{-(d_{ji}/\alpha_H)^2}}{C_H}\left(1-\frac{r_0H(i,t)}{k}\right)\exp\left[r_0\left(1-e^{-Q(i,t)/\delta}-\frac{H(i,t)}{k}\right)\right]\exp\left(-\frac{aP(i,t)}{1+awP(i,t)}\right)\right\}\\ &+\sum_{j=1}^{n}\lambda_P(j,t+1)\left\{\frac{e^{-(d_{ji}/\alpha_P)^2}}{C_P}\left[1-\exp\left(-\frac{aP(i,t)}{1+awP(i,t)}\right)\right]\right\}-\lambda_Q(i,t)\frac{uv}{[v+H(i,t)]^2},\\ \lambda_P(i,t) = \frac{\partial\mathcal{H}}{\partial P(i,t)} = &-\sum_{j=1}^{n}\lambda_H(j,t+1)\left\{\frac{e^{-(d_{ji}/\alpha_H)^2}}{C_H}H(i,t)\exp\left[r_0\left(1-e^{-Q(i,t)/\delta}-\frac{H(i,t)}{k}\right)\right]\frac{a}{[1+awP(i,t)]^2}\exp\left(-\frac{aP(i,t)}{1+awP(i,t)}\right)\right\}\\ &+\sum_{j=1}^{n}\lambda_P(j,t+1)\left\{\frac{e^{-(d_{ji}/\alpha_P)^2}}{C_P}H(i,t)\frac{a}{[1+awP(i,t)]^2}\exp\left(-\frac{aP(i,t)}{1+awP(i,t)}\right)\right\},\\ \lambda_Q(i,t) = \frac{\partial\mathcal{H}}{\partial Q(i,t)} = &\sum_{j=1}^{n}\lambda_H(j,t+1)\left\{\frac{e^{-(d_{ji}/\alpha_H)^2}}{C_H}H(i,t)\frac{r_0}{\delta}e^{-Q(i,t)/\delta}\exp\left[r_0\left(1-e^{-Q(i,t)/\delta}-\frac{H(i,t)}{k}\right)\right]\exp\left(-\frac{aP(i,t)}{1+awP(i,t)}\right)\right\}\\ &+\lambda_Q(i,t+1)\beta+W(i,t),\end{aligned} \tag{S30}$$

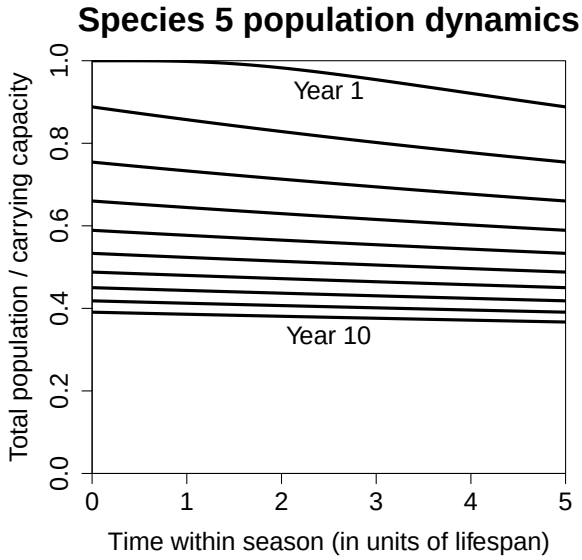
with terminal conditions

$$\lambda_H(i,T) = \lambda_P(i,T) = 0, \quad \lambda_Q(i,T) = W(i,T) \quad \text{for all } i.$$

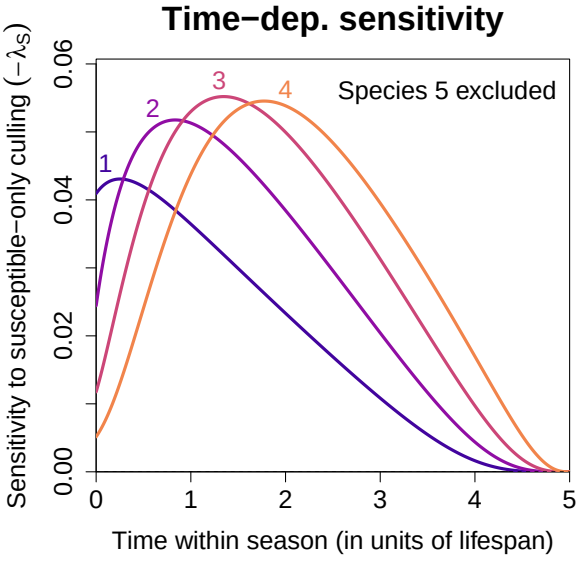
# S7 Supplementary figures and tables from Example 1: Exogenous disease spillover in multi-species sink networks



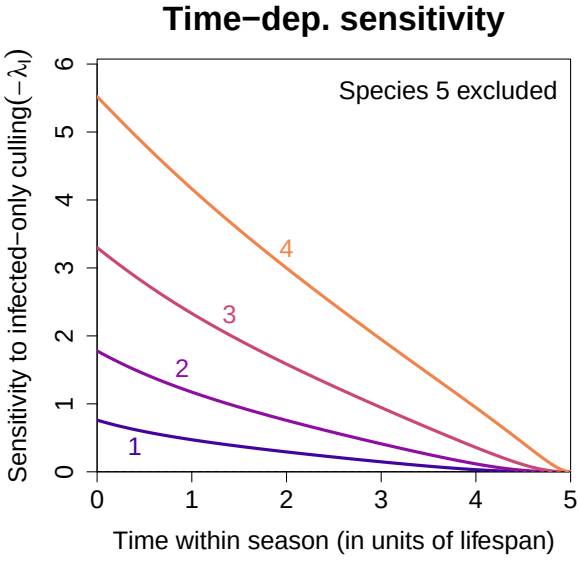

(A)




(B)

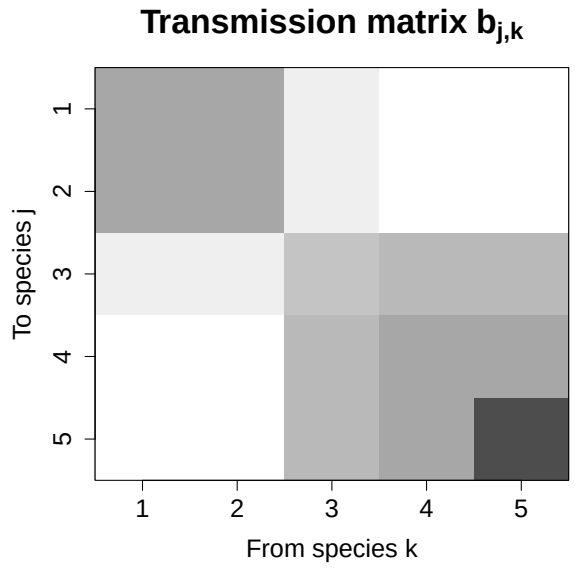



(C)

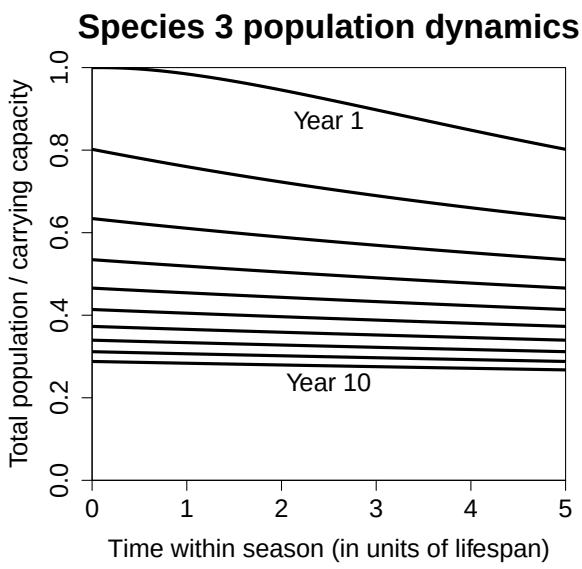



(D)

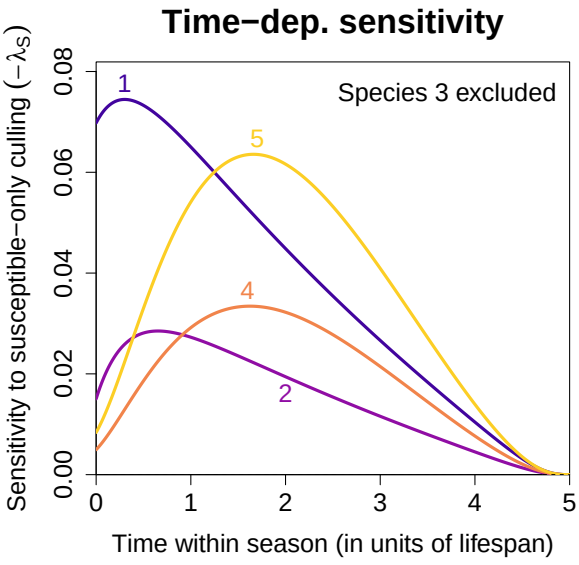
Figure S1: Additional figures from Network 1. **(A)** Matrix representation of the transmission coefficients $b_{j,k}$. **(B)** Population decline in the species of concern (species 5) over a 10-year period, assuming that the population size at the end of one season carries over to the start of the next season. The purpose is to show that the population decline can be significant despite the low infection prevalence shown in Fig. 3(D). **(C)** Time-dependent sensitivity when only susceptible individuals are culled. **(D)** Time-dependent sensitivity when only infected individuals are culled $(-\lambda_{I_j})$. The weighted sum of (C) and (D) gives the time-dependent sensitivity to indiscriminate culling $(-\lambda_{N_j})$ shown in Fig. 3(G).



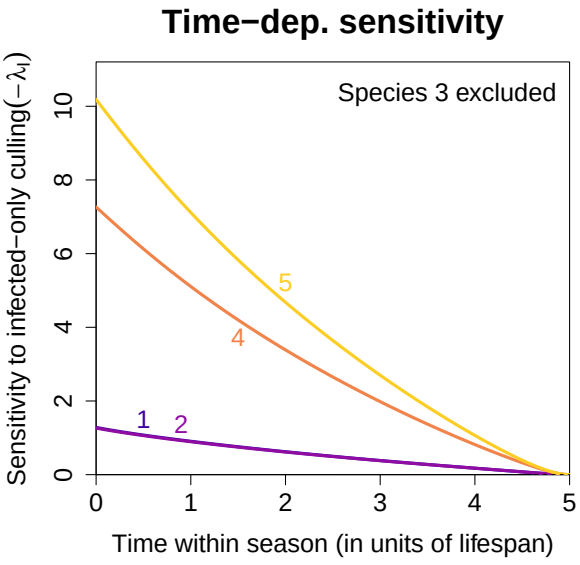

(A)




(B)

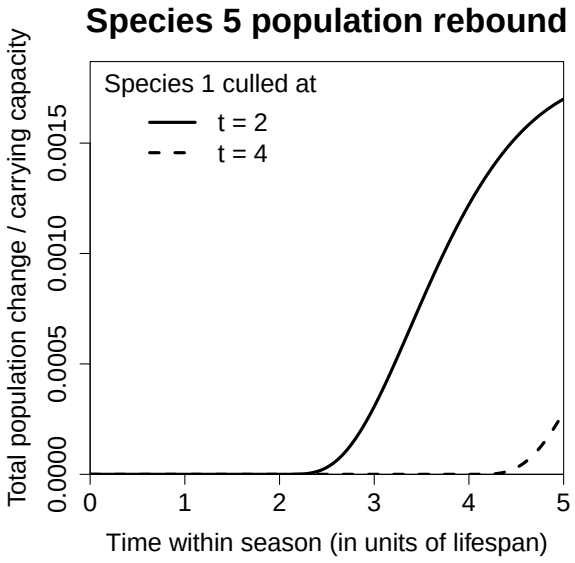



(C)

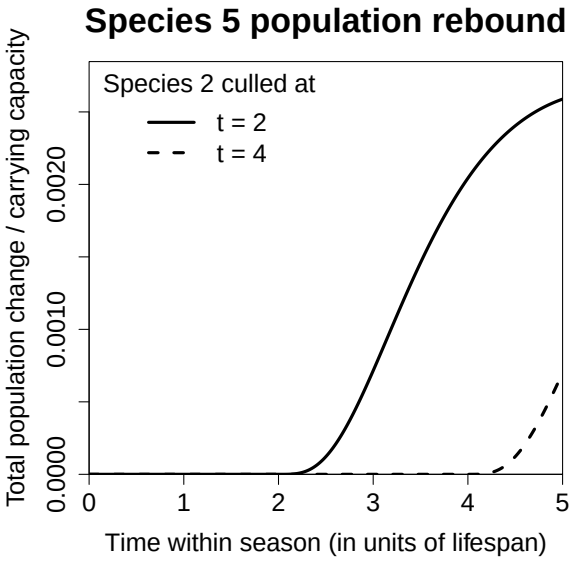



(D)

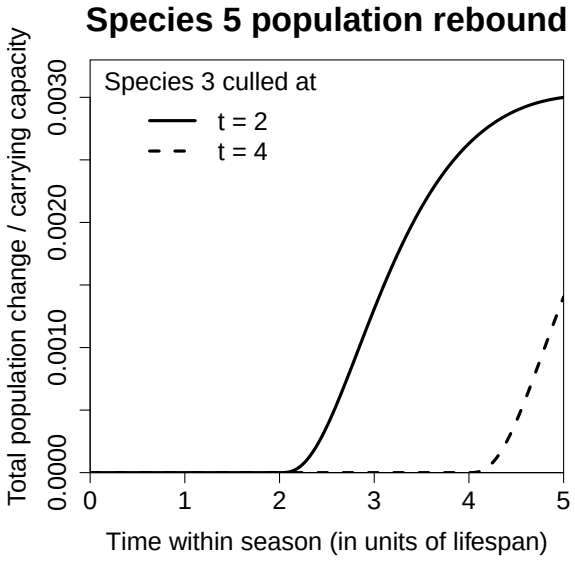
Figure S2: Similar to Fig. S1, except for Network 2.



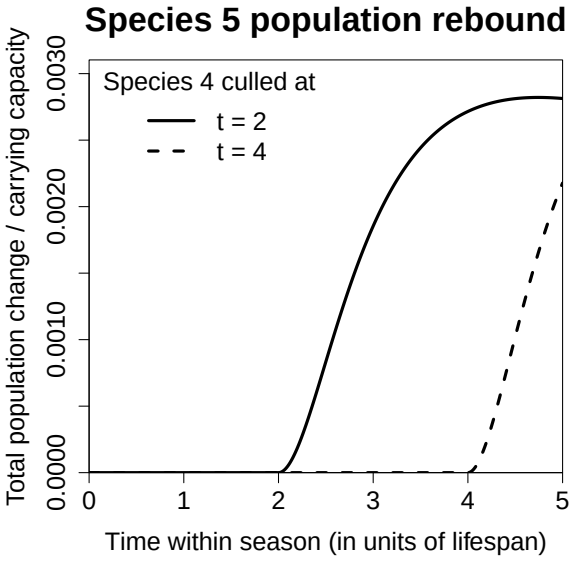

(A)




(B)

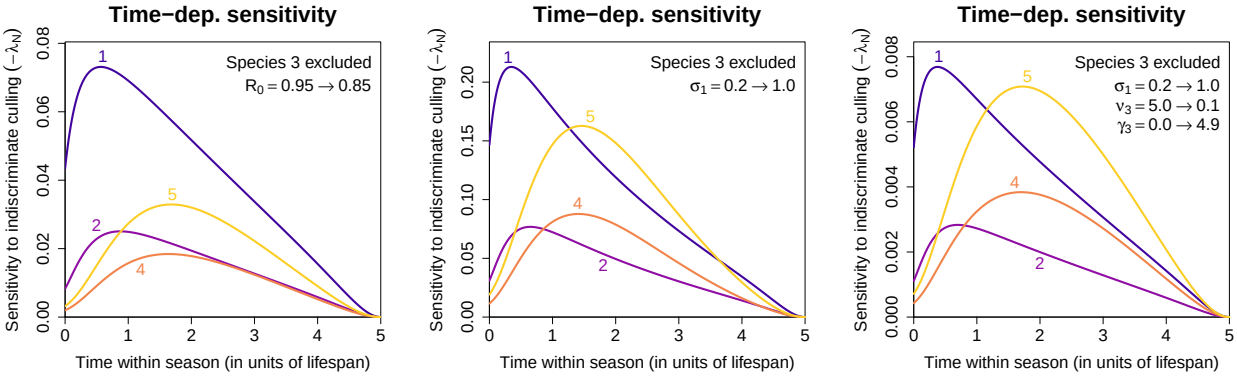



(C)




(D)

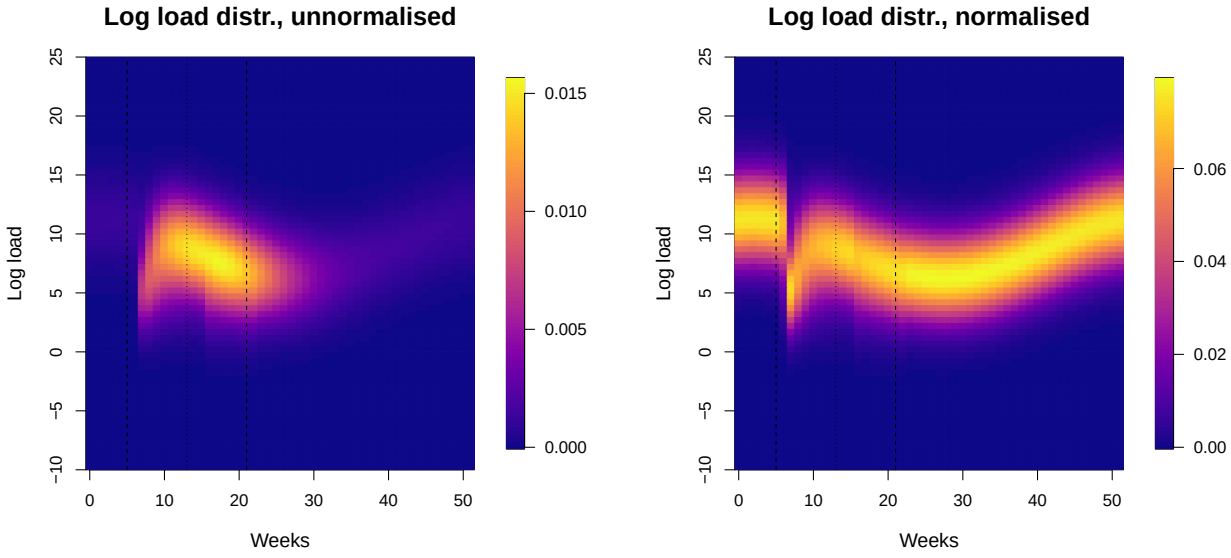
Figure S3: For Network 1, the graphs above show the population rebound in the species of concern (species 5) when 10% of another species is indiscriminately culled. Late culling leaves less time for the population to rebound (affecting the terminal payoffs $V_{S_{j_C}}$ and $V_{I_{j_C}}$), and also less time for the rebound to contribute to the integral in the reward function.




(A) (B) (C)

Figure S4: More results from Network 2, obtained using modified parameter values. **(A)** Reducing $R_0$ caused the importance of species 5 to fall entirely below species 1, due to multi-step within-module transmission becoming less likely at a lower $R_0$. **(B)** Increasing the exogenous spillover rate $\sigma_1$ caused the most important species to switch from species 5 back to species 1 towards the end of the season. This is due to the large decrease in the population of species 3 resulting from the increased spillover; the switch no longer occurred in **(C)** when we converted most of the disease-induced mortality rate in species 3 to its recovery rate.

# S8 Supplementary figures and tables from Example 2: Leopard frogs as reservoirs of the amphibian chytrid fungus




(A) (B)

Figure S5: **(A)** Number of infected frogs in each log load bin, each week across the year, at steady state. **(B)** Log load distribution each week, obtained by normalizing the sum of each vertical column in (A) to 1. Due to the temperature-dependent load dynamics, we see that the load is the lowest in summer and the highest in winter.

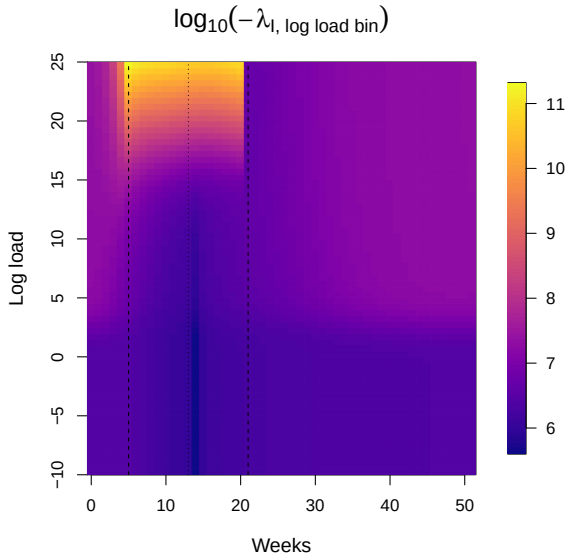



Figure S6: The sensitivity to removing an infected frog from each log load bin, each week across the year. Note that this sensitivity does not take into account whether the log load bin is actually "occupied" which is why we choose to work with $-\lambda_I(t)$ as defined in Eqn. (24) instead.

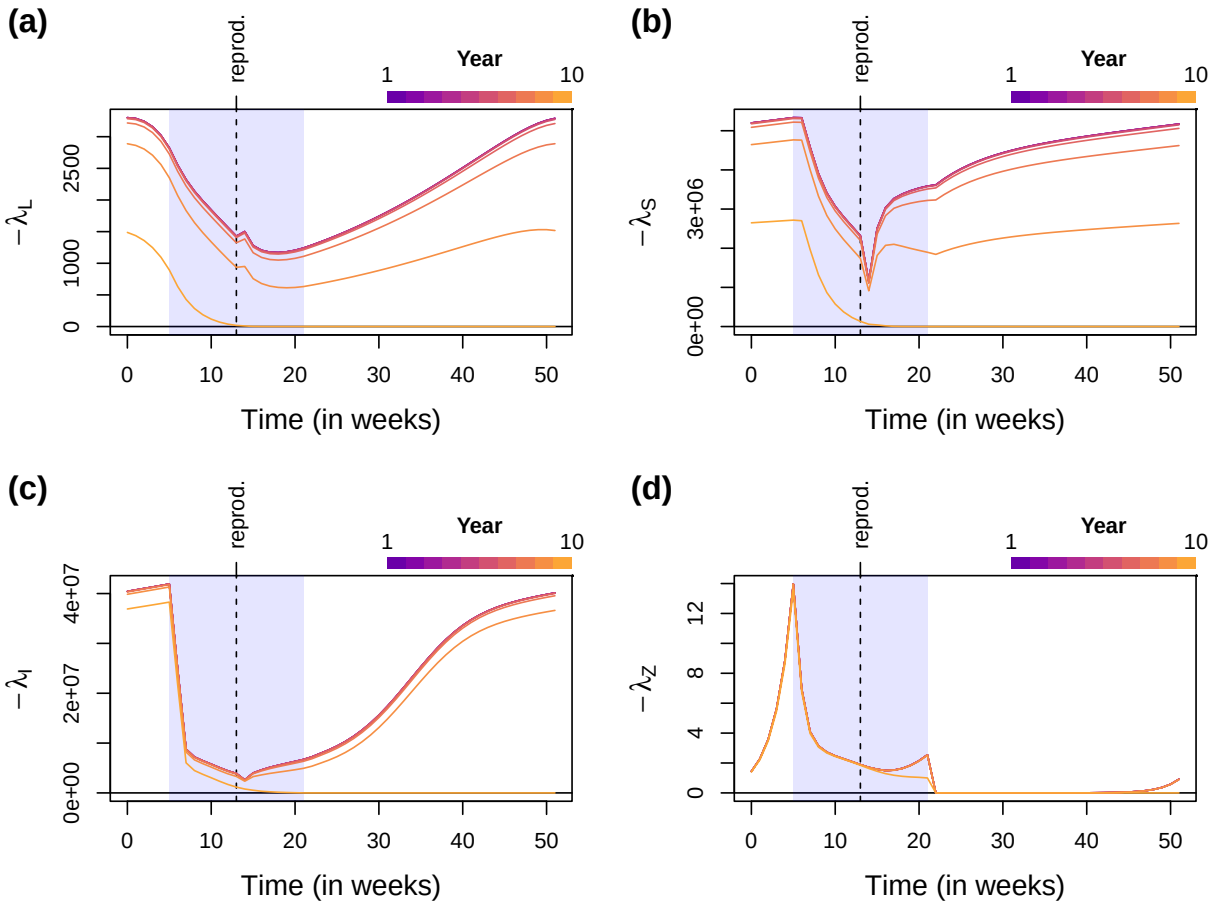



Figure S7: **Effects of the time horizon $T$.** Similar to Fig. 6, except that we have also shown the sensitivities every year within the time horizon. We see that if the time horizon is sufficiently long, the seasonal sensitivity patterns during the first few years are identical. At steady state, each year starts with the same "initial conditions", so the second year can be thought of as the same system with a time horizon of 9 years, the third year a time horizon of 8 years, etc. Therefore, the fact that the early years show identical seasonal patterns means that the early-year patterns are independent of the time horizon, and hence expected to be the same as when the time horizon is infinite.

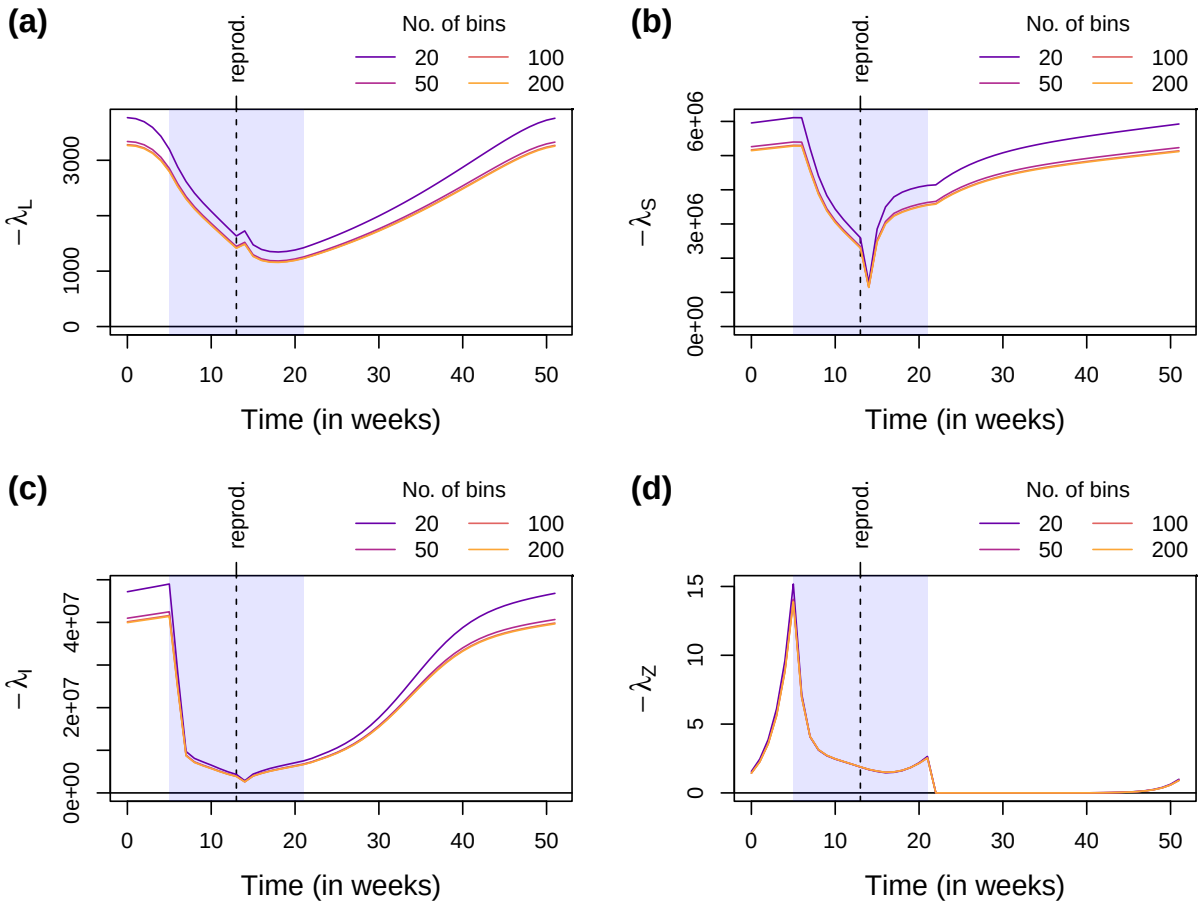



Figure S8: **Varying the number of bins in the discretized IPM.** Similar to Fig. 6, except that we have varied the number of bins used when discretizing the IPM.

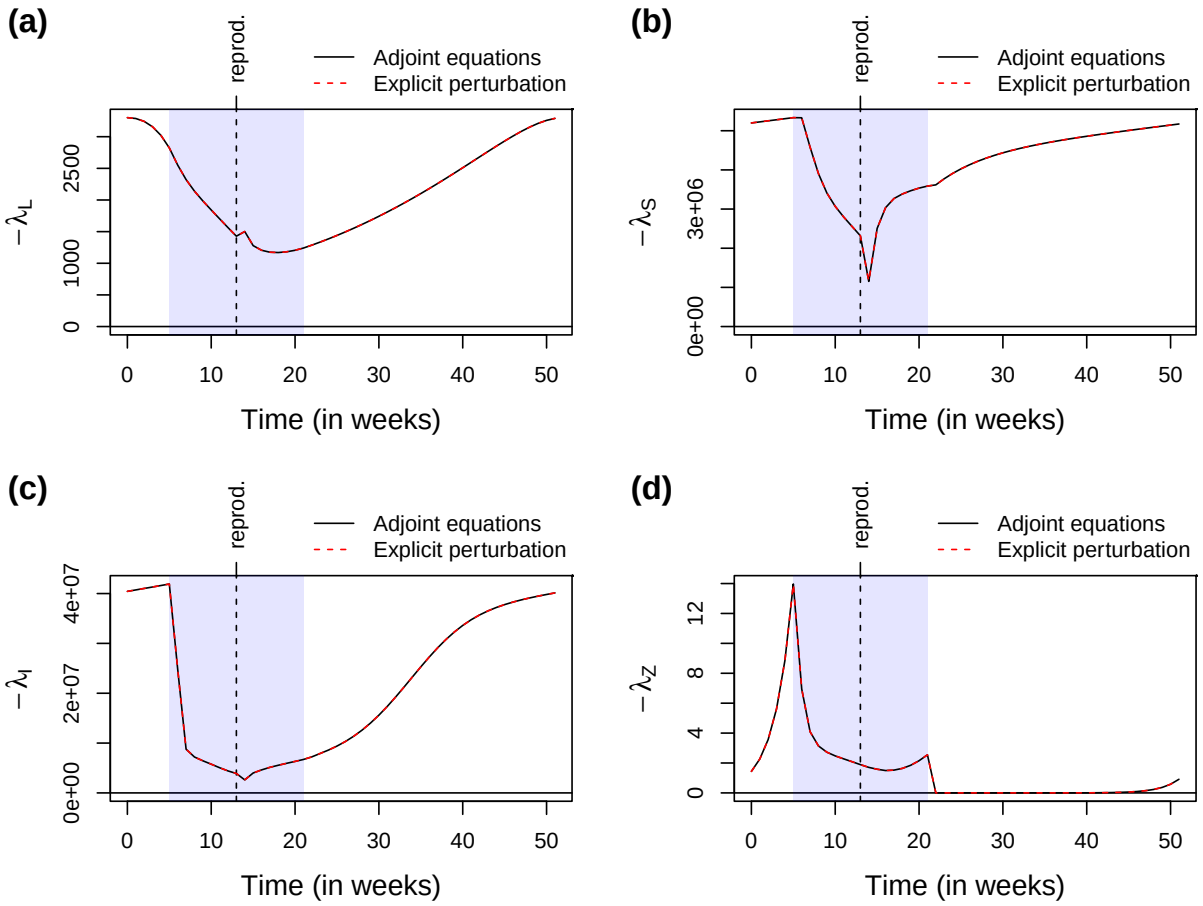


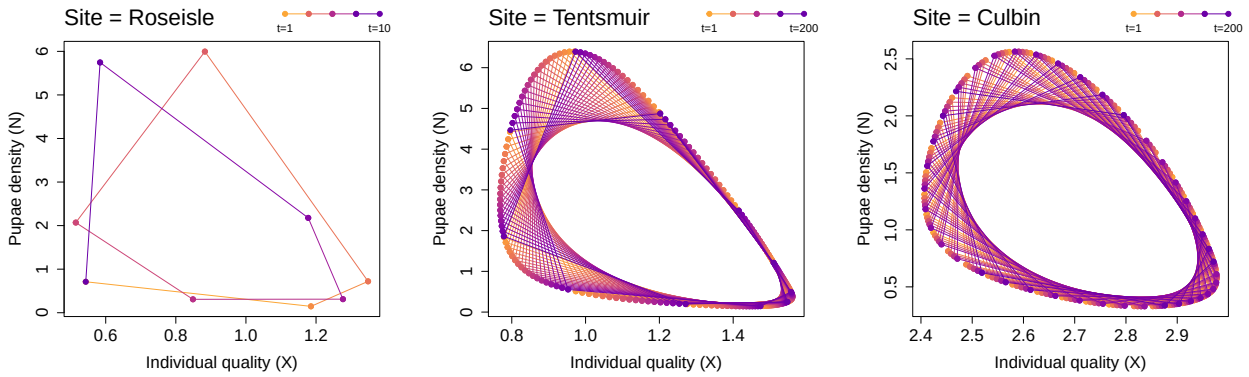

Figure S9: **Checking against explicit perturbations.** Similar to Fig. 6, except that we have also shown the sensitivities obtained by explicitly perturbing the state variables at each time point (red dashed lines). The perfect agreement with the adjoint variables implies that the adjoint equations have been correctly derived and implemented.

# S9 Supplementary figures and tables from Example 3: Population cycles in the pine looper and the larch budmoth

## *S9.1 Pine looper*

| Site | $r$ | $s$ | $u$ | $x_{\min}$ | $\beta$ |
|---|---|---|---|---|---|
| Culbin | $5.064\times10^{-5}$ | 0.079 | 3.364 | 2.150 | 0.204 |
| Roseisle | $5.760\times10^{-2}$ | 0.246 | 3.644 | 0.510 | 1.016 |
| Tentsmuir | $5.677\times10^{-3}$ | 0.000 | 4.075 | 0.618 | 0.294 |

Table S1: Parameter values of the maternal effects model, fitted separately using data at three sites.




(A) (B) (C)

Figure S10: Phase plane diagram at Roseisle, Tentsmuir and Culbin, showing the periodic steady-state solution at Roseisle, and the quasiperiodic steady-state solutions at Tentsmuir and Culbin. At Roseisle, we only showed 10 years to illustrate one complete cycle of two oscillations, whereas at Tentsmuir and Culbin, we showed every year across the time horizon of 200 years.

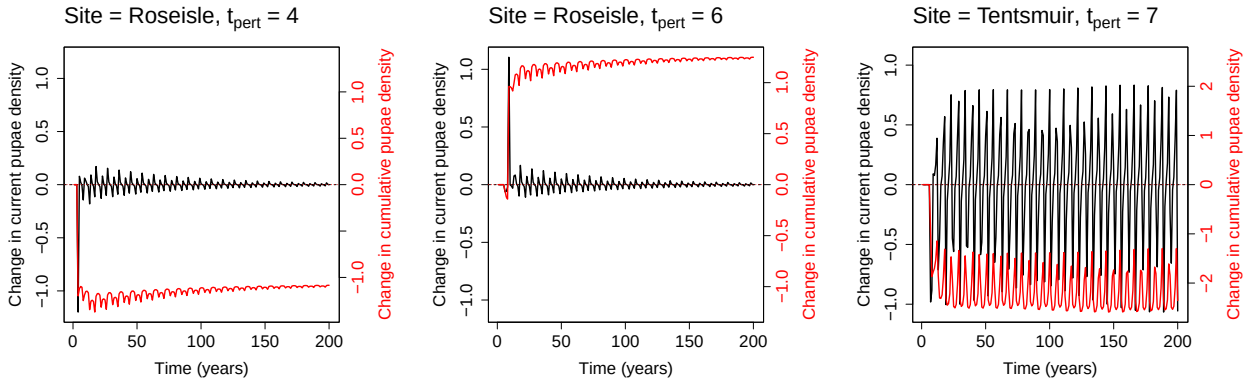



Figure S11: Changes in the current pupae density $N(t)$ and the cumulative moth density $\sum_{t'=1}^{t} N(t')$ at all $t$, following a 20% cull at $t = t_{\text{pert}}$. **(A)** Roseisle; $t_{\text{pert}} = 4$. **(B)** Roseisle; $t_{\text{pert}} = 6$. **(C)** Tentsmuir; $t_{\text{pert}} = 7$. We see that the changes in current density decay with time in (A) and (B), but persist indefinitely in (C), likely because of the steady-state trajectories being periodic in Roseisle, but quasiperiodic in Tentsmuir. As a result, the cumulative changes approach constant, non-oscillatory values in (A) and (B), but remain oscillatory in (C). Note that the choices of $t_{\text{pert}}$ are unimportant here; we made these specific choices only to facilitate comparison with Fig. 7(D-F) and Fig. S12.

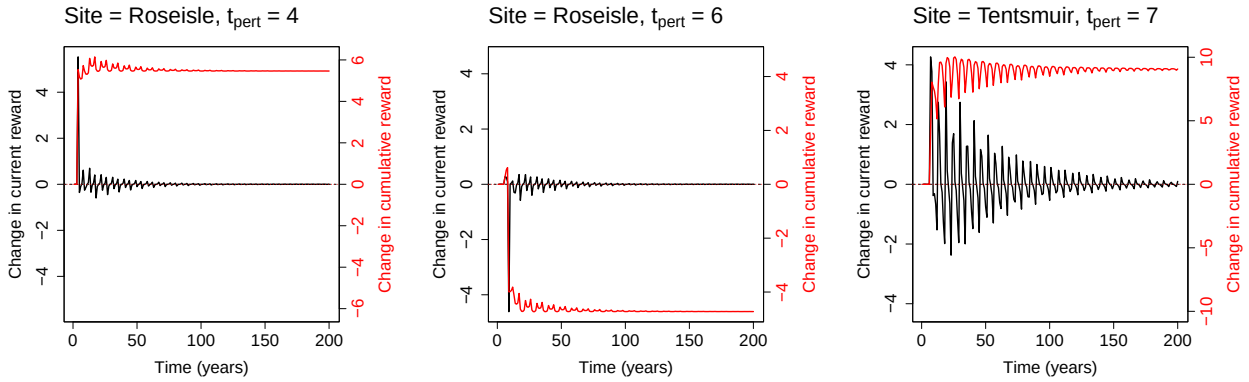
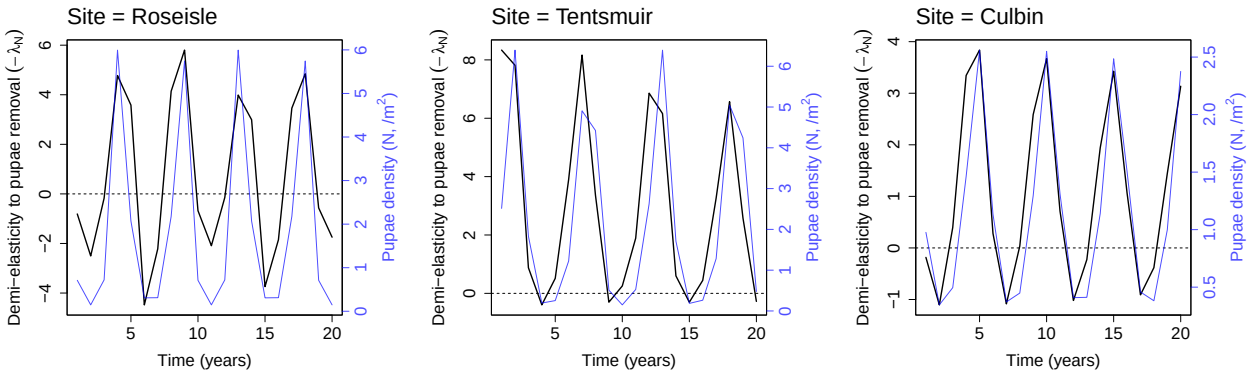


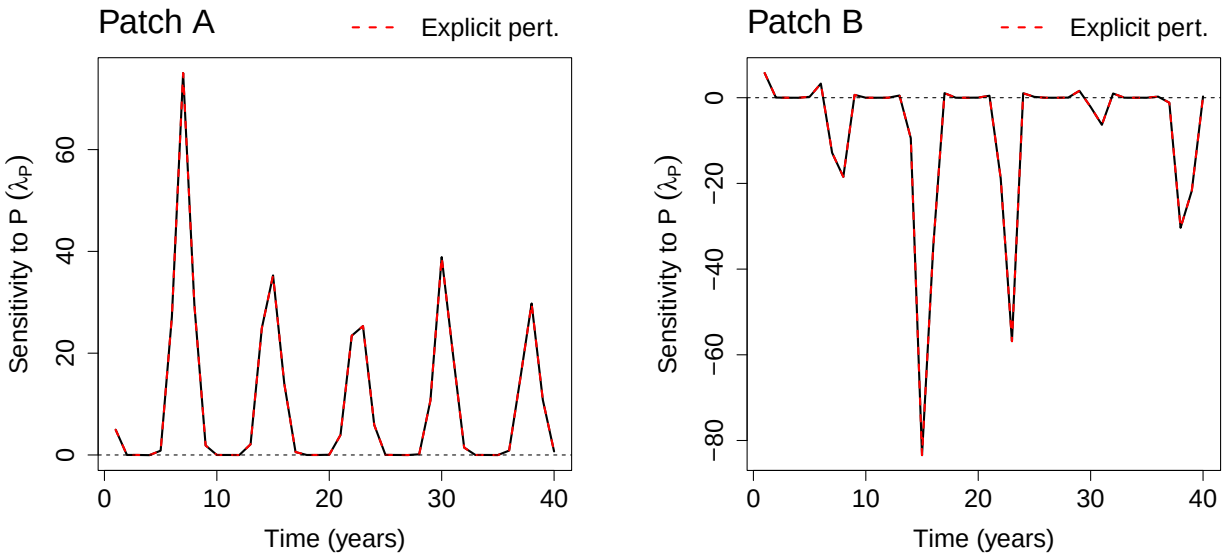

Figure S12: Changes in the current reward $-N(t)W(t)$ and the cumulative reward $-\sum_{t'=1}^{t} N(t')W(t')$ at all $t$, following a 20% cull at $t = t_{\text{pert}}$. We have rescaled these changes by a factor of $1/0.2$, so that the cumulative reward at $t = T = 200$ should be approximately equal to the demi-elasticity in Fig. S13 at $t = t_{\text{pert}}$; any small discrepancies are due to nonlinearities from the relatively large perturbation. **(A)** Roseisle; $t_{\text{pert}} = 4$. **(B)** Roseisle; $t_{\text{pert}} = 6$. **(C)** Tentsmuir; $t_{\text{pert}} = 7$. Note that unlike Fig. S11(C), the changes in current reward decay in time because of the decaying weight $W(t)$. This allows the cumulative reward to approach a constant, non-oscillatory value, and hence reduces the dependence of the demi-elasticities on the time horizon $T$.




(A) (B) (C)

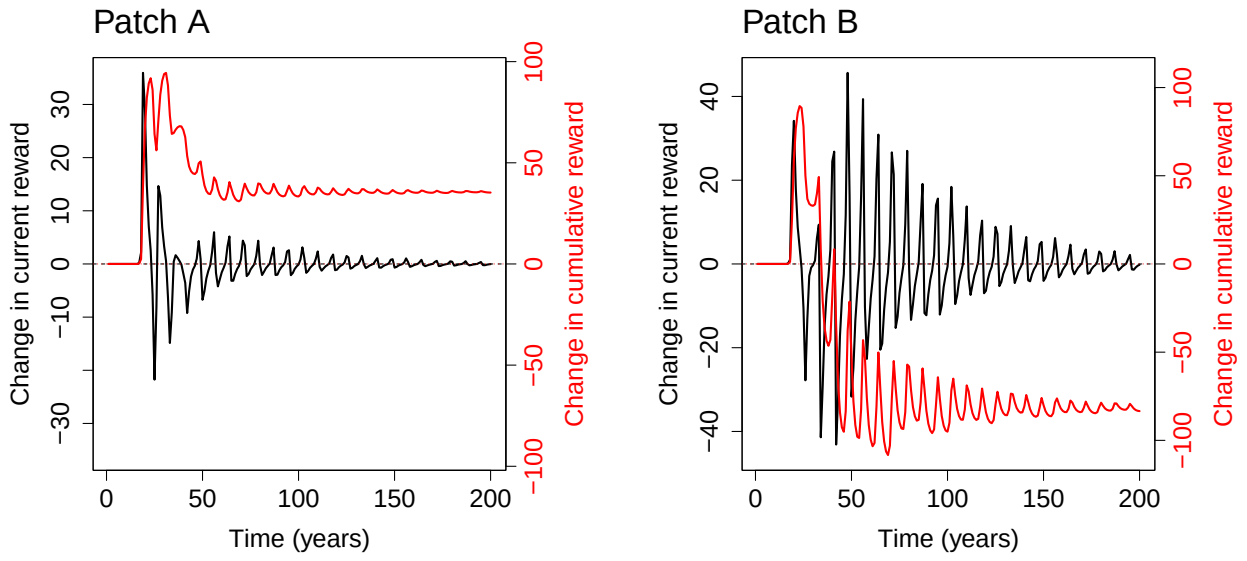
Figure S13: Demi-elasticities of the reward to the culling of pine looper at **(A)** Roseisle, **(B)** Tentsmuir and **(C)** Culbin.

## *S9.2 Larch budmoth*




(A) (B)

Figure S14: Verifying that TDSA gives the correct sensitivities for the larch budmoth model using explicit perturbations. We focused on the two patches discussed in Fig. 8.




(A) (B)

Figure S15: The effects of adding parasitoids at $t=15$ to the two patches discussed in Fig. 8. The current reward refers to the sum of plant quality times the weight in the current year, and the cumulative reward the sum of current rewards from $t=1$ up to the current year. We used small perturbations to ensure linearity, but scaled the results by the inverse of the perturbation size, so that the change in cumulative reward at $t=T=200$ (the end of the time horizon) should be equal to the sensitivity at $t=15$ (the time of perturbation). As expected, they indeed agree with Fig. S14 at $t=15$ ($\sim 40$ for Patch A, $\sim -80$ for Patch B).